



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
[www.uspto.gov](http://www.uspto.gov)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/818,607	03/28/2001	Yasuo Okutani	35.G2761	1901
5514	7590	05/11/2006	EXAMINER	
FITZPATRICK CELLA HARPER & SCINTO 30 ROCKEFELLER PLAZA NEW YORK, NY 10112			WOZNIAK, JAMES S	
			ART UNIT	PAPER NUMBER
			2626	

DATE MAILED: 05/11/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

<b>Office Action Summary</b>	Application No.	Applicant(s)
	09/818,607	OKUTANI ET AL.
	Examiner James S. Wozniak	Art Unit 2626

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --  
**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

#### Status

- 1) Responsive to communication(s) filed on 22 February 2006.
- 2a) This action is FINAL.                    2b) This action is non-final.
- 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

#### Disposition of Claims

- 4) Claim(s) 1,6,7,9-13,18,19 and 21-33 is/are pending in the application.
  - 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) Claim(s) \_\_\_\_\_ is/are allowed.
- 6) Claim(s) 1,6,7,9-13,18,19 and 21-33 is/are rejected.
- 7) Claim(s) \_\_\_\_\_ is/are objected to.
- 8) Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

#### Application Papers

- 9) The specification is objected to by the Examiner.
- 10) The drawing(s) filed on 3/28/2001 is/are: a) accepted or b) objected to by the Examiner.  
 Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
 Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

#### Priority under 35 U.S.C. § 119

- 12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
  - a) All    b) Some \* c) None of:
    1. Certified copies of the priority documents have been received.
    2. Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
    3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

#### Attachment(s)

- |   |   |
|---|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892)   | 4) <input type="checkbox"/> Interview Summary (PTO-413)                     |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)                                    | Paper No(s)/Mail Date. _____.   |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)<br>Paper No(s)/Mail Date _____. | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152) |
|   | 6) <input type="checkbox"/> Other: _____.                                   |

## DETAILED ACTION

### *Response to Amendment*

1. In response to the office action from 11/22/2005, the applicant has submitted an amendment, filed 2/22/2006, amending claims 1, 9, 13, 21, 25, and 28, while adding claims 32-33 and arguing to traverse the art rejection based on the amended limitations (*Amendment, Pages 12-20*). The applicant's arguments have been fully considered but are moot with respect to the new grounds of rejection in view of Donovan et al ("The IBM Trainable Speech Synthesis System," 1998).

### *Claim Rejections - 35 USC § 112*

2. The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

3. **Claims 1, 6-7, 9-12 and 32** are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

In claim 1, applicants claimed:

- Means for modifying each of the synthesized units according to prosody information based on the input text; and

- Means for selecting synthesis units based on the modification distortions obtained by said distortion obtaining means.

In claim 32, applicants claimed:

- Means for obtaining a respective modification distortion for each of the plurality of synthesis units from a modification distortions table according to prosody information obtained based on the input text; and
- Means for selecting synthesis units based on the modification distortions obtained by said distortion obtaining means.

With respect to Claim 1, in regards to the means for modifying synthesis units with respect to an input text prosody, the specification shows the selection of a set of synthesis unit candidates based on an input text prosody (*Pages 7-8*), but fails to show a structure of a means for modifying the candidate units based on an input text prosody. Rather, the modification distortion appears to be the result of PSOLA processing (*Pages 1 and 13-14*) and the specification does not show a structure of a means for such processing utilizing the prosody of an input text.

With respect to Claim 32, in regards to the means for obtaining a modification distortion for synthesis units with respect to an input text prosody, the specification shows the storage of modification distortions in a table (*Pages 19-20*), but fails to show a structure of a means for retrieving the modification distortions for synthesis units from the table based on an input text prosody.

With respect to claims 1 and 32, in regards to the means and step for selecting synthesis units based on the modification distortions, the specification shows the selection of a synthesis

Art Unit: 2626

unit based on a joint a modification and concatenation distortion (Pages 14-15), but fails to show a structure of a means for selecting a synthesis unit based solely on a modification distortion.

Thus, the specification does not disclose adequate structure for performing the recited functions, thereby failing to particularly point out and distinctly claim the invention as required by the second paragraph of section 112. Because no structure disclosed in the embodiments of the invention actually performs the claimed functions, the specification lacks the corresponding structure as required by 35 U.S.C. 112, 6<sup>th</sup> paragraph, and fails to comply with 35 U.S.C. 112, 2<sup>nd</sup> paragraph.

*"If one employs means plus function language in a claim, one must set forth in the specification an adequate disclosure showing what is meant by that language. If an applicant fails to set forth an adequate disclosure, the applicant has in effect failed to particularly point out and distinctly claim the invention as required by the second paragraph of section 112." In re Donaldson Co., 16 F.3d 1189, 1195, 29 USPQ2d 1845, 1850 (Fed. Cir. 1994) (in banc).*

*If there is no disclosure of structure, material or acts for performing the recited function, the claim fails to satisfy the requirements of 35 U.S.C.112, second paragraph. Budde v. Harley-Davidson, Inc., 250 F.3d 1369, 1376, 58 USPQ2d 1801, 1806 (Fed. Cir. 2001); Cardiac Pacemakers, Inc. v. St. Jude Med., Inc., 296 F.3d 1106, 1115-18, 63 USPQ2d 1725, 1731-34 (Fed. Cir. 2002). MPEP 2100-217.*

The written description is objected to in light of 35 U.S.C. 112 1<sup>st</sup> paragraph for failing to show any corresponding structure of the claimed means for obtaining a modification distortion and selecting a synthesis unit based on a modification unit. (*See In re Knowlton, 481 F.2d 1357, 1366, 178 USPQ 486, 492–93 (CCPA 1973). Conversely, the invocation of 35 U.S.C. 112, sixth*

*paragraph does not exempt an applicant from compliance with 35 U.S.C. 112, first and second paragraphs. See Donaldson, 16 F.3d at 1195, 29 USPQ2d at 1850; Knowlton, 481 F.2d at 1366, 178 USPQ at 493. See MPEP 2100-217-218.)*

4. The following is a quotation of the first paragraph of 35 U.S.C. 112:

The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

5. **Claims 1, 6-7, 9-13, 18-19, 21-33** are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the enablement requirement. The claim(s) contains subject matter that was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention.

With respect to Claims 1, 13, and 25, the specification recites an apparatus and method that selects of a set of synthesis unit candidates based on an input text prosody (Pages 7-8), but fails to disclose how a synthesis unit is modified according to prosody of an input text to obtain a modification distortion.

With respect to Claims 32 and 33, the specification mentions the storage of modification distortions in a table (Pages 19-20), but fails teach how a specific modification distortion can be obtained from such a table based on prosody information found in text.

With respect to claims 1, 13, 25, and 32-33, the specification teaches the selection of a synthesis unit based on a joint a modification and concatenation distortion (Pages 14-15), but fails to recite how a synthesis unit is selected solely based on a modification distortion.

Dependent claims 6-7, 9-12, 18-19, 21-24, and 26-31 do not remedy the lack of

enablement issue noted above with respect to claims 1, 13, 25 and 32-33, and therefore, are also rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the enablement requirement.

***Claim Rejections - 35 USC § 103***

6. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

7. **Claims 1, 12-13, 24-25, and 31** are rejected under 35 U.S.C. 103(a) as being unpatentable over Eide et al (*U.S. Patent: 6,101,470*) in view of Donovan et al ("The IBM Trainable Speech Synthesis System," 1998).

With respect to **Claim 1, 13, and 25**, Eide recites:

Obtaining means for obtaining a plurality of synthesis units based on an input text (*obtaining phonemes corresponding to an input text, Col. 3, Lines 35-49; and Col. 4, Lines 32-40*);

Modifying means for modifying each of the synthesis units according to prosody information obtained based on the input text (*applying stress levels to a phoneme sequence, Col. 4, Lines 41-60*);

A selection means for selecting synthesis units based on a distance measure (Col. 8, Lines 42-53); and

Speech synthesis means for performing speech synthesis based on the synthesis units selected by said selection means (Col. 3, Lines 45-49).

Eide further teaches method implementation as a program stored on a computer readable medium (*Col. 2, Line 64- Col. 3, Line 34*).

Although Eide teaches a means for selecting synthesis units based on a distance measure, the distance measure utilized by Eide does not involve a distortion based on synthesis units before and after modification. Donovan, however, teaches a pitch modification cost that would effectively measure the difference between an original synthesis unit and a pitch modified synthesis unit (*cost of pitch modifying a synthesis unit, Pages 2-3, Section 4.1*).

Eide and Donovan are analogous art because they are from a similar field of endeavor in text-to-speech synthesis. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings of Eide with the pitch modification cost taught by Donovan in order to provide a further means to ensure high quality synthetic speech (*Page 1, Abstract*).

With respect to **Claims 12, 24, and 31**, Eide recites:

Input means for inputting text data (*Col. 3, Lines 35-49*);

Language analysis means for performing language analysis of the text data (Col. 4, Lines 32-40);

Prosody-parameter generation means for generating predetermined prosody parameters based on a result of analysis of said language analysis means (*Col. 3, Lines 35-49*);

Wherein said distortion obtaining means obtains the modification distortion based on the predetermined prosody parameters generated by said prosody parameter generation means (*Col. 4, Lines 9-26; and Col. 8, Lines 42-53*).

8. **Claims 6-7, 18-19, and 26-27** are rejected under 35 U.S.C. 103(a) as being unpatentable over Eide et al in view of Donovan et al, and further in view of Huang et al (*U.S. Patent: 5,913,193*).

With respect to **Claims 6, 18, and 26**, Eide in view of Donovan teaches the speech synthesis apparatus and method that utilizes a pitch modification distortion cost, in selecting a best speech unit for synthesizing speech, as applied to Claims 1, 13, and 25. Eide in view of Donovan does not teach obtaining a distortion by adding modification and concatenation distortion, however Huang discloses:

A speech signal processing apparatus and method, wherein the distortion obtaining means uses a value obtained by adding the obtained modification distortion between the synthesis units before and after modification and a concatenation distortion (*spectral distortion between adjacent instances, Col. 3, Lines 1-6*) generated by concatenating a synthesis unit to another synthesis unit (*summing the distortions of an instance sequence, Col. 9, Lines 44-47*).

Eide, Donovan, and Huang are analogous art because they are from a similar field of endeavor in speech synthesis. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings of Eide in view of Donovan with the method of summing distortions including a concatenation distortion as taught by Huang in order

to provide more natural synthesized speech generation by minimizing spectral distortion between speech segment boundaries (Huang, Col. 1, Line 57- Col. 2, Line 9).

With respect to **Claims 7, 19, and 27**, Huang recites:

A speech signal processing apparatus and method, wherein the distortion obtaining means calculates a weighted sum of the modification distortion between the synthesis units before and after modification and the concatenation distortion generated by concatenating a synthesis unit to another synthesis unit (*Col. 8, Line 51- Col. 9, Line 22*).

9. **Claims 9, 21, and 28** are rejected under 35 U.S.C. 103(a) as being unpatentable over Eide et al view of Donovan et al, and further in view of Akamine et al (*U.S. Patent: 6,161,091*).

With respect to **Claims 9, 21, and 28**, Eide in view of Donovan teaches the speech synthesis apparatus and method that utilizes a pitch modification distortion cost, in selecting a best speech unit for synthesizing speech, as applied to Claims 1, 13, and 25. Eide in view of Donovan does not specifically suggest calculating modification distortion using a cepstrum distance, however Akamine discloses:

A speech signal processing apparatus and method, wherein said distortion obtaining means calculates the modification distortion using a cepstrum distance (*Col. 5, Line 56- Col. 6, Line 19*).

Eide, Donovan, and Akamine are analogous art because they are from a similar field of endeavor in speech synthesis. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings of Eide in view of Donovan with the means of calculating distortion through cepstral distance as taught by Akamine in order to

provide a well-known means that better describes speech segments, in addition to Euclidean distance, for determining a most accurate phonetic segment for the generation of more natural synthesized speech (Akamine, *Col. 5, Line 56- Col. 6, Line 19; and Col. 4, Lines 27-30*).

10. **Claims 10-11, 22-23, 29-30, and 32-33** are rejected under 35 U.S.C. 103(a) as being unpatentable over Eide et al in view of Donovan et al, and further in view of Coorman et al (*U.S. Patent: 6,665,641*).

With respect to **Claims 10, 22, 29, 32, and 33**, Eide in view of Donovan teaches the speech synthesis apparatus and method that utilizes a pitch modification distortion cost, in selecting a best speech unit for synthesizing speech, as applied to Claims 1, 13, and 25. Eide in view of Donovan does not teach the use of a table to determine a modification distortion, however Coorman discloses:

A speech signal processing apparatus and method, wherein the distortion obtaining means includes a table storing distortions, and determines the modification distortion by referring to the table (*Col. 13, Line 33- Col. 14, Line 21*).

Eide, Donovan, and Coorman are analogous art because they are from a similar field of endeavor in speech synthesis. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings Eide in view of Donovan with the use of a table for determining a modification distortion as taught by Coorman in order to provide a means for easily selecting candidate speech units that most closely match target speech (*Coorman, Col. 9, Lines 27-38*).

With respect to **Claims 11, 23, and 30**, Eide in view of Donovan teaches the speech synthesis apparatus and method that utilizes a pitch modification distortion cost, in selecting a best speech unit for synthesizing speech, as applied to Claims 1, 13, and 25. Eide in view of Donovan does not teach the use of a table to determine a concatenation distortion, however Coorman discloses:

A speech signal processing apparatus and method, wherein the distortion obtaining means includes a table storing distortions, and determines the concatenation distortion by referring to the table (*Col. 11, Lines 43-67; Col. 14, Lines 23-49, and Col. 7, Lines 43-50*).

Eide, Donovan, and Coorman are analogous art because they are from a similar field of endeavor in speech synthesis. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings Eide in view of Donovan with the use of a table for determining concatenation distortion as taught by Coorman in order to provide a means for easily selecting candidate speech units that will not cause pitch discontinuities (*Coorman, Col. 9, Lines 39-44*).

### ***Conclusion***

11. Any inquiry concerning this communication or earlier communications from the examiner should be directed to James S. Wozniak whose telephone number is (571) 272-7632. The examiner can normally be reached on M-Th, 7:30-5:00, F, 7:30-4, Off Alternate Fridays.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David Hudspeth can be reached at (571) 272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

James S. Wozniak  
4/20/2006



DAVID HUDSPETH  
SUPervisory Patent Examiner  
Technology Center 2600